A Nonlinear Krylov Accelerator for the Boltzmann k-Eigenvalue Problem

Matthew T. Calef, Erin D. Fichtl, James S. Warsa, Markus Berndt, Neil N. Carlson, CCS-2

We compare variants of Anderson Mixing with the Jacobian-Free Newton-Krylov and Broyden methods applied to the k-eigenvalue formulation of the linear Boltzmann transport equation. We present evidence that one variant of Anderson Mixing finds solutions in the fewest number of iterations. We examine and strengthen theoretical results of Anderson Mixing applied to linear problems.

The criticality of fissioning systems is typically described by the k-eigenvalue formulation of the linear Boltzmann transport equation. Physically, the largest eigenvalue, generally denoted by k, is the effective neutron multiplication factor that, in the equation, scales the fission production term to achieve a steady-state solution. The corresponding eigenmode describes the neutron flux profile for that steady-state (i.e., critical) system and, when the system is close to criticality, provides useful information about the distribution of the neutron population in space and velocity.

Mathematically, the equation is a standard eigenproblem for which power iteration is well-suited because the eigenmode of interest is most commonly that with the largest magnitude. For the k-eigenvalue problem, each step of a true power iteration incurs a heavy computational cost due to the expense of fully inverting the transport operator, therefore a nonlinear fixed point iteration (FPI) is generally employed in which an approximate inversion of this operator is performed at each iteration. Recently, more sophisticated nonlinear iteration methods than FPI, predominately Jacobian-Free Newton-Krylov (JFNK) , have been applied with great success [1,2].

No broad comparison of the performance of nonlinear solvers applied to the k-eigenvalue problem has been performed. In [1] we present such a comparison, examining the performance of three nonlinear solvers, JFNK, Broyden's Method, and Anderson Mixing, applied to the k-eigenvalue problem. A variant of Anderson Mixing [2], first described in [3], is of particular interest because, in the experience of the authors, it is frequently computationally more efficient than JFNK and Broyden's Method.

JFNK is an inexact Newton's method in which the inversion of the Jacobian is performed to arbitrary precision using a Krylov method (most commonly the generalized minimum residual method, GMRES) and the Jacobian itself is never formed, but rather its action approximated using finite differences of arbitrarily close state data. JFNK can be expected to converge quadratically in a neighborhood containing the solution. Each iteration of JFNK requires a nested inner iteration and the bulk of the computational effort is expended in this inner Krylov inversion of the Jacobian at each outer Newton step. At the end of each inversion, the accumulated Krylov space is discarded even though the Jacobian is expected to change minimally during the final Newton steps when a similar space will be rebuilt in the next Newton iteration. In effect, at the end of each iteration, JFNK discards information that may be of use in successive iterations.

In its standard formulation, Broyden's method uses differences in state from successive iterations to make low-rank updates to the Jacobian. The Sherman-Morrison-Woodbury update rule is then used to compute the action of the inverse of the Jacobian after such an update. While Broyden's method is restricted to low-rank updates, it provides an explicit representation of the Jacobian, allowing one to employ the Dennis-More condition to show that it converges super-linearly in a neighborhood containing the solution.

Anderson Mixing [2] uses differences in state from successive iterations to infer information about the inverse of the Jacobian, which is assumed to be roughly constant in a neighborhood containing all the iterates. Unlike Broyden's method, the updates can be of arbitrary rank. Recent results by Walker and Ni [6] show that, with mild assumptions, Anderson Mixing applied to a linear problem performs as well as GMRES. In this regard, Anderson Mixing may be thought of as a nonlinear version of GMRES. In independent work, Carlson and Miller formulated a so-called nonlinear Krylov acceleration [3] method, which we show to be a variant of Anderson Mixing. We refer to this method as NKA.

Our test problem is a cylinder with a 3.5-cm radius and a height of 9 cm modeled in 2D cylindrical coordinates. The problem consists of a central 5-cm layer of Boron-10 with 1-cm-thick water layers on either side and 1-cm layers of highly enriched uranium on the ends. The top, bottom,

For more information contact Markus Berndt at berndt@lanl.gov

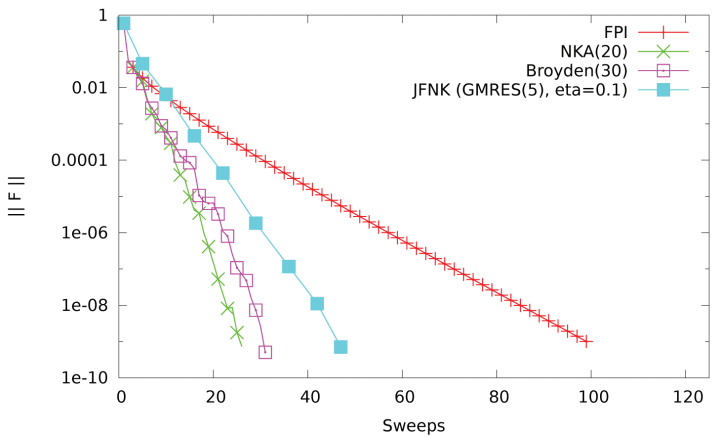

Fig. 1. The PARTISN code applied to the unreflected cylinder problem. Comparison of the performance of four nonlinear solvers. Scaled L2-norm of the residual as a function of number of sweeps for the various methods and subspace sizes. Note that, for JFNK, points plotted on the lines indicate when a JFNK iteration starts (JFNK requires multiple sweeps per iteration). For the other methods, we plot one point per two iterations of the method.

and radial boundaries are either all vacuum or all reflective (the inner radial boundary is a symmetry condition in cylindrical coordinates). This a difficult problem (i.e., one with a high dominance ratio) because the problem is symmetric, having two fissile regions that are effectively decoupled because of the reflector (water) and absorber (boron) between them, which means the second eigenmode is close to the fundamental mode–reflection further increases the difficulty of the problem. A 16-group Hansen-Roach cross-section data set is used to generate the results presented here.

A 175 (r-axis) by 450 (z-axis) mesh of equally sized squares is used in PARTISN [7] for both the reflected and unreflected problems. The Capsaicin results are computed on an unstructured mesh comprising roughly the same number of cells in the r and z axes as the PARTISN computation, for a total of 79,855 (possibly non-convex) polygons, each of which has 3 to 6 vertexes on a cell. While these two codes are both deterministic, discrete ordinates transport codes, they are quite different in the spatial discretizations they employ and in other aspects of implementation, which can cause differences in iterative behavior and result in slightly different solutions.

In Fig. 1, we present a comparison of methods described above for the unreflected cylinder problem run with PARTISN. The particular runs that are compared here represent the best set of parameters that we found for each of the methods. Points plotted on the lines represent two nonlinear iterations of the method, except for JFNK where we plot a point at the beginning of each JFNK inner iteration. We plot the error on the vertical axis, while on the horizontal axis we plot the number of sweeps that, for this class of problems, is an excellent measure of computational cost. The results shown in Fig. 1 are representative of our more detailed study [1]. NKA performs best, with Broyden second, JFNK third; FPI is not competitive. Moreover, NKA exhibited the least sensitivity with respect to its parameters.

The benefit of JFNK is its Newton-like convergence, achieved by developing an arbitrarily accurate approximation of the inverse of the Jacobian at each of these outer iteration, but the cost is repeated function evaluations–or sweeps in our case–in each of these outer iterations and the wasteful discarding of potentially useful subspace information. In contrast, Broyden and Anderson Mixing only perform a single function evaluation at each iteration, but continue to use old information from previous iterations to improve their estimate of the Jacobian. The drawback for these methods is that the approximate Jacobian or its inverse is based on an amalgam of new and old information, so they are unlikely to converge in fewer iterations than Newton's method. Performance of all these methods will clearly depend on how the Jacobian is changing from iteration to iteration, and how information collected at each function evaluation is used. Memory requirements for Broyden are half that of NKA, making it an attractive alternative.

Our numerical results indicate that Anderson Mixing in the form of NKA found solutions in the PARTISN and Capsaicin codes for both the reflected and unreflected problems in the fewest number of function evaluations and the shortest runtimes. These physics problems represent actual reactor geometries and were run at a large scale. Our results highlight the strength of this method: regularity, consistency, and efficiency. In our results, NKA was shown to bring the norm down to zero smoothly, much as FPI and JFNK do, but with greater efficiency than those methods. Broyden at times achieved excellent performance, did not always demonstrate this same smooth convergence behavior, and often diverged. Based on these results we feel that NKA may be well-suited to other computational physics problems.

[1] Calef, M.T. et al., *J Comput Phys,* submitted (2012).
[2] Anderson, D.G., *J ACM* **12**, 547 (1965).
[3] Carlson, N.N. and K. Miller, *SIAM J Sci Comput* **19**, 728 (1998).
[4] Gill, D.F. et al., *Nucl Sci Eng* **168,** 37 (2011).
[5] Knoll, D.A. et al., *Nucl Sci Eng* **167,** 133 (2011).
[6] Walker, H.F. and P. Ni, *SIAM J Numer Anal* **49,** 1715 (2011).
[7] Alcouffe, R.E. et al., LANL Technical Report LA-UR-05-3925 (2005).

Funding Acknowledgments
DOE, NNSA Advanced Simulation and Computing Program